Translation and Stationarity for Graph Signals

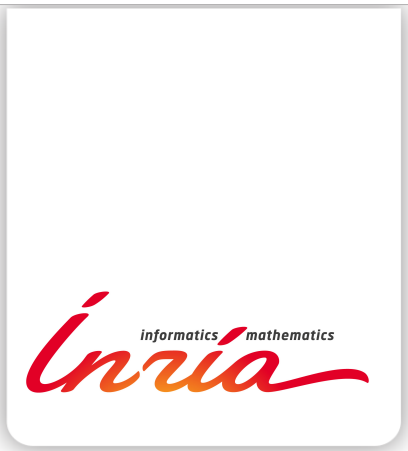
Benjamin Girault, Paulo Gonçalves, Eric Fleury

Translation and Stationarity for Graph Signals

Benjamin Girault, Paulo Gonçalves, Éric Fleury

RESEARCH REPORT
N° 8719
April 2015
Project-Teams DANTE

ISSN 0249-6399 ISRN INRIA/RR--8719--FR+ENG



Translation and Stationarity for Graph Signals

Benjamin Girault*, Paulo Gonçalves†, Éric Fleury*†

Project-Teams DANTE

Research Report n° 8719 — April 2015 — 18 pages

Abstract: We propose a novel time shift like operator called the *graph translation*. We enforce that it is isometric and define it such that it shares with the time shift operator its key properties. Using the graph translation operator and its isometry, we propose a tractable definition of stationary graph signals and characterise stationary graph signals using their spectral properties. We illustrate this characterisation on synthetic stationary random graph signals. Finally we propose a method to test the stationarity of a real dataset.

Key-words: Graph signal processing, graph translation, stationarity

* École Normale Supérieure de Lyon
† Inria Rhône-Alpes

RESEARCH CENTRE
GRENOBLE – RHÔNE-ALPES

Inovallée
655 avenue de l'Europe Montbonnot
38334 Saint Ismier Cedex

Translation et stationnarité pour les signaux sur graphe

Résumé : Nous proposons un nouvel opérateur similaire à l'opérateur de translation en temps que nous appelons *graph translation*. Nous forçons cet opérateur à être isométrique et le définissons tel qu'il partage un maximum de propriétés avec l'opérateur de translation en temps. La propriété d'isométrie de notre opérateur nous permet alors de proposer une définition de signaux stationnaires sur graphe et une caractérisation spectrale de ces signaux. Nous illustrons alors cette caractérisation sur des signaux aléatoires synthétiques et stationnaires. Enfin, nous proposons une méthode pour évaluer de manière empirique la stationnarité des signaux présent dans des données de terrain.

Mots-clés : Traitement du signal sur graphe, translation sur graphe, stationnarité

Contents

1 Introduction

Following the increasing quality and quantity of data collected, we see more and more interesting datasets that need to be interpreted. In particular, many of these datasets are structured into entities linked together. Thus they form a graph. In turn these entities support data. Using tools borrowed and adapted from the classical signal processing toolbox, the emerging field of graph signal processing aims at studying these structured datasets. While classical signal processing has been studying signals supported by well defined Euclidean structures, general graphs face the challenge of not having the combinatorial comfort of the properties of Euclidean structures.

Yet, the field has seen several successes recently [11], and the toolbox of graph signal processing is growing stronger. Among the challenges not yet fully addressed, the problem of an equivalence of the time shift operator is of particular interest. Indeed, the time shift operator is at the very core of temporal signal processing, from the action of "time passing by", to linear time shift invariant systems, or stationary signals defined as the time-shift invariance of their stochastic properties.

Stationarity is the property of being independent from the time origin. This property for temporal signals have been study at length over the years and stationary time signals are well understood. Moreover, the vast majority of interesting temporal signal are non-stationary, meaning that the mathematically tractable property of stationarity is essential to study the interesting non-stationary signals. We propose here a tractable definition of stationary graph signals, and use it on synthetic data and real data.

2 Background

2.1 Time Series

Let $x[n] = x(n\tau_s)$ be a time series with τ_s as sampling period. Let T be the *time shift* operator such that $T\{x\}[n] = x[n-1]$. This operator is at the very core of time series since its action is that of "time passing by". Also, the entire field of *Linear Time Shift Invariant* (LTSI) filtering is dedicated to operators invariant by this time shift. Among its important properties, this operator is linear, acts as a convolution of the signal with a Kronecker's delta δ_1 located at $n = 1$, and is isometric: $\|T\{x\}\|_2 = \|x\|_2 = (\sum_n |x[n]|^2)^{1/2}$.

Another central tool for time series analysis is the Fourier transform given by $X(\omega) = \sum_n x[n]e^{-\imath\omega n\tau_s}$, for any (angular) frequency ω. By duality of the time sampling, $X(\omega)$ is periodic of period $\frac{2\pi}{\tau_s}$, thus allowing to restrict the study of $X(\omega)$ to $[0, \frac{2\pi}{\tau_s}]$. For simplicity, reduced frequencies $\widetilde{\omega} = \omega\tau_s$ are used in the definition of the Discrete Fourier Transform (DFT) and its inverse:

$$X(\widetilde{\omega}) = \sum_{n=-\infty}^{+\infty} x[n]e^{-\imath\widetilde{\omega}n}, \quad \widetilde{\omega} \in [0, 2\pi] \tag{1}$$

$$x[n] = \frac{1}{2\pi}\int_0^{2\pi} X(\widetilde{\omega})e^{\imath\widetilde{\omega}n}d\widetilde{\omega}. \tag{2}$$

In the analysis of N-periodic time series, *i.e.* $(N\tau_s)$-periodic signals, the set of reduced frequencies $\omega_k = 2\pi\frac{k}{N}$ with $k \in \{0, \dots, N-1\}$ is a basis of this space of periodic time series. Moreover, (2) becomes:

$$x[n] = \frac{1}{N}\sum_{k=0}^{N-1} X(\omega_k)e^{\imath\omega_k n}. \tag{3}$$

In the DFT, lower reduced frequencies are close to 0 or 2π, and higher frequencies close to π. For example, the time series $x[n] = \cos(\frac{1}{N}n)$ gives $X(\omega_1) = X(\omega_{N-1}) = \frac{N-1}{2}$ and $X(\omega_k) = 0$ elsewhere, coherent with x being a low frequency time series.

Finally, the eigenvectors of the time shift are the Fourier modes $e_\omega[n] = e^{\imath\omega n}$, and its eigenvalues the complex exponentials $e^{-\imath\omega}$ such that:

$$T\{e_\omega\}[n] = e^{-\imath\omega} e_\omega[n]. \tag{4}$$

Therefore, the time shift behaves as a phase shifting operator. In this letter, we propose a new shifting operator on graph signals by analogy to these properties.

2.2 Signal Processing on Graphs

We start by presenting the signal processing on graphs, the novel field studying signals carried by vertices of a graph. Let $\mathcal{G} = (V, E)$ be a graph with V the set of vertices and $E \subseteq V \times V$ the set of edges between vertices. Let $N = |V|$ be the number of vertices. The adjacency matrix of the graph $\mathcal{G}$ is A with a_{ij} the weight of the edge ij or 0 if no edge links i to j. A graph signal is a function $X : V \to \mathbb{C}$ assigning a sample to each vertex. X is represented as a column vector of $\mathbb{C}^N$ with its i^{th} component being the sample on the i^{th} vertex.

We consider in this letter only symmetric graphs for which the adjacency matrix is symmetric. We define the Laplacian matrix $L = D - A$ where D is the diagonal matrix of degrees $d_{ii} = \sum_j a_{ij} = d_i$. L being symmetric positive semidefinite [3], it has a set of orthonormal eigenvectors χ_l associated with non-negative eigenvalues λ_l, with $L\chi_l = \lambda_l\chi_l$ and $\lambda_0 \leq \lambda_1 \leq \ldots \lambda_{N-1}$. We will suppose the graph to be connected, *i.e.* there is a path between any two pair of vertices in the graph, in which case L has only one eigenvalue zero [3].

Performing signal processing on such a graph consists in finding a meaningful set of modes to analyse signals. Most often [11], the positive semidefinite property of the Laplacian is used to define the Fourier transform matrix F as the projection on the basis of the eigenvectors of L $\{\chi_0, \ldots, \chi_{N-1}\}$. The same property ensures that $F^{-1} = F^*$ and an equivalence of Parseval's identity is verified. The Laplacian matrix is then $L = F^*\Lambda F$, with Λ the diagonal matrix of the eigenvalues λ_l.

We denote $\widehat{X} = FX$ the projection of the signal X on the Fourier basis. The (generalised) convolution of two signals is then defined as a multiplication in the Fourier domain [13]:

$$(X * Y)(i) = \sum_{l=0}^{N-1} \widehat{X}(l)\widehat{Y}(l)\chi_l(i). \tag{5}$$

The associated convolutive operator is $H = F^* \operatorname{diag}(\widehat{X})F$ such that $HY = X*Y$. More generally, given a linear operator H, $\widehat{H} = FHF^*$ denotes the expression of H in the Fourier domain.

Given the complexity of the graph structure, and the properties it misses compared to Euclidean spaces, a definition of signal processing on graphs that everyone would agree on is challenging. In particular, several key tools of classical signal processing have been recently generalised to the graph structure using different approaches. Examples include the Fourier transform [11, 10], wavelet decomposition [4, 6], or the very basic operation of translation that is at the core of temporal signal processing [10, 12]. The translation being the subject of this letter, we present two translation operators recently proposed.

2.2.1 Generalised Translations

D. Shuman, B. Ricaud & P. Vandergheynst define the *generalised translations* in [13] from the observation that shifting a time series by k samples is equivalent to performing the convolution

of the signal with a Kronecker's delta located at k. A generalised translation operator is then an operator T_k such that $T_k X = \sqrt{N}\,\delta_k * X$. In the Fourier domain, this can be written as $\widehat{T_k} = \sqrt{N}\,\mathrm{diag}(\widehat{\delta_k})$, *i.e.* as a diagonal matrix of diagonal $\sqrt{N}\widehat{\delta_k}$. These operators are linear and convolutive.

2.2.2 Graph Shift

A. Sandryhaila and J. Moura proposed in [10] to use the linear operator defined by the adjacency matrix A as the equivalent of the time shift for graph signals, and call it the *graph shift*. They define then the Fourier transform as the projection on the eigenvectors of the graph shift. This definition is possible for any kind of graph, symmetric or not, and any weight on the edges. However, if A is not diagonalisable, then the Jordan decomposition is used to obtain generalised eigenvectors. Using the algebraic signal processing framework [8], the authors define an equivalence of LTSI filters, called the *Linear Graph Shift Invariant* (LGSI) filters. The graph shift is justified by analogy to the signal processing on periodic time series. Indeed, the periodic time series can be seen as signals on a circular oriented graph with as many vertices as samples in a period. In this context, the graph shift is exactly the time shift for N-periodic time series. In this letter, we will only consider symmetric graphs such that A will always be diagonalisable. The graph shift is then an operator *shifting* the value on one vertex to its neighbours according to the weights of the edges.

3 Graph Translation

Both the generalised translations and the graph shift are not isometric operators with respect to the ℓ_2-norm. This property being mathematically comfortable to obtain a tractable invariance property for graph signals, we propose in this section to define a new time shift-like operator verifying it. Also, similarly to the time shift, the graph shift and the generalised translations, we define a graph translation operator that is both linear and convolutive. We will denote $T_{\mathcal{G}}$ the matrix representation this operator and call it the *graph translation*. Since $T_{\mathcal{G}}$ is a convolutive operator, $\widehat{T_{\mathcal{G}}}$ is a diagonal matrix.

3.1 Enforcing Isometry

We design the graph translation as an *isometric* operator with respect to the ℓ_2-norm $\|X\|_2 = (\sum_i |X(i)|^2)^{1/2}$. From an energy point of view, this is equivalent to saying that the energy of the signal is left unchanged. The use of this norm follows the convention of graph signal energy found in the literature [10, 12]. Since $\widehat{T_{\mathcal{G}}}$ is diagonal, isometry forces each component of this diagonal to be unimodular. Therefore, $\widehat{T_{\mathcal{G}}} = \exp(\imath\widehat{\Omega})$ with $\widehat{\Omega}$ a diagonal matrix, leading to:

$$T_{\mathcal{G}} = \exp(\imath\Omega). \tag{6}$$

3.2 The matrix Ω

Given the general form (6), we now need to specify the matrix Ω to identify a shift operator. We propose to proceed by analogy to the temporal case to specify the expression of $T_{\mathcal{G}}$. Indeed, in the context of time series, the phase shift of a mode of frequency ω by the time shift operator is $-\omega$ as shown in (4). We build upon this observation to define $\widehat{\Omega}$.

To define frequencies on a graph, we follow the observation in [6] in which the graph Laplacian L can be seen as an approximation of the continuous Laplacian Δ (up to a change of sign). More

precisely, the (temporal) Fourier modes being eigenvectors of the continuous Laplacian:

$$\Delta e^{\imath\omega t} = -\omega^2 e^{\imath\omega t}, \tag{7}$$

the eigenvalues of the continuous Laplacian are the opposite of the squared frequencies. In other words, the phase shifts are the square roots of the eigenvalues of the continuous Laplacian. This parallel can also be found in [1] where the authors prove that the graph Laplacian is an approximation of the Laplace-Beltrami operator on a manifold and that the graph Laplacian converges to the Laplace-Beltrami operator as the sampling of the manifold densifies.

Back to our graph setting, we propose to transpose this into the *graph frequencies* $\omega_l = \sqrt{\lambda_l}$, $-\omega_l$ being then the l^{th} entry of the matrix $\widehat{\Omega}$ by analogy to (4):

$$\widehat{\Omega} = -\sqrt{\Lambda}, \tag{8}$$

with Λ the non-negative diagonal matrix of the eigenvalues of L. Choosing between positive and negative signs for the graph frequencies ω_l is an open question. In the temporal case, part of the answer lies in the reduced frequencies definition. However, another part of the answer is in the sign of frequencies: positive for $[0, \pi]$, negative for $(\pi, 2\pi) = (-\pi, 0) \mod 2\pi$. For graphs, frequencies are in general not paired by opposite frequencies, and choosing one sign or the other is a question that still needs to be elucidated. We propose in this letter to use only positive frequencies since the eigenvalues conceptually represent the spatial span of the energy of the Fourier modes. The question of reduced frequencies is addressed in the next section.

3.3 Reduced Graph Frequencies

We saw that for time series, the Fourier transform is studied on the interval $[0, \frac{2\pi}{\tau_s}]$. This interval is rescaled to match the interval $[0, 2\pi]$, defining the reduced frequencies. In this section, we propose a definition of *reduced frequency* for graphs such that they lie in the interval $[0, \pi]$, according to our choice of considering only positive frequencies.

To achieve this goal, we propose to use theoretical bounds on λ_{N-1} proposed in the literature. The authors of [5] have proven the following upper bound (among others) on the eigenvalues of the Laplacian matrix:

$$\rho_{\mathcal{G}} = \max_{i \in V} \sqrt{2d_i(d_i + \bar{d}_i)} \quad \text{with} \quad \bar{d}_i = \frac{\sum_{ij \in E} w_{ij} d_j}{d_i}. \tag{9}$$

This upper bound has the interesting property of being met only when the graph is *bipartite regular*, in other words when the vertices of the graph can be partitioned in two sets with edges only between those two sets (bipartite), and all vertices have same degree $d_i = d$ (regular). In particular, the bound is met for circular graphs with even numbers of vertices, and not met for an odd number of vertices. This is similar to the DFT where the reduced frequency $\omega_k = \pm\pi$ is met only when N is even, in which case $k = \frac{N}{2}$.

Since we have a bound on the eigenvalues of the Laplacian, we can rescale those eigenvalues such that the graph frequencies lie in the interval $[0, \pi]$:

$$\widetilde{\lambda}_l = \pi^2 \frac{\lambda_l}{\rho_{\mathcal{G}}}. \tag{10}$$

We remark that rescaling all eigenvalues of the Laplacian by the same factor is equivalent to rescaling the weights of the graph by the same factor: $U(\alpha\Lambda)U^* = \alpha U\Lambda U^* = \alpha D - \alpha A$, where $\alpha = \pi^2/\rho_{\mathcal{G}}$. This is similar to time series where frequencies are rescaled such that they come down to the case $\tau_s = 1$. The following definition introduces then the notion of reduced graph frequency.

Definition 1 (Reduced Graph Frequency)**.** *The* reduced frequency *associated to the Fourier mode χ_l is:*

$$\widetilde{\omega_l} = \pi\sqrt{\lambda_l/\rho_{\mathcal{G}}}$$

Using this definition, we can now introduce our graph translation operator verifying $\widehat{T_{\mathcal{G}}} = \mathrm{diag}(e^{-\imath\widetilde{\omega_0}}, \ldots, e^{-\imath\widetilde{\omega_{N-1}}})$ and such that the translated $T_{\mathcal{G}}X$ of X reads:

Definition 2 (Graph Translation)**.**

$$T_{\mathcal{G}} = \exp\left(-\imath\pi\sqrt{L/\rho_{\mathcal{G}}}\right).$$

3.4 Graph Translation Properties

To illustrate the complexity of designing an isometric operator, we consider the simple operation of normalising the output of an operator H. This composed operator produces an isometric operator H^{iso}. However, unless $X \mapsto \|HX\|_2$ is a constant function, it is not linear, nor does it preserve the *Power Spectral Density* (PSD) $|\widehat{X}|^2$ of the signal. A better approach to define H^{iso} is to enforce $H^{\mathrm{iso}} = \exp(\imath\Omega)$. If $\widehat{H}$ is diagonal, *i.e.* H is a convolutive operator, and $|\widehat{H}| \leq 1$, then we can enforce $\mathrm{Re}(H^{\mathrm{iso}}) = H$ by using $\Omega = \mathrm{acos}(H)$: $H^{\mathrm{iso}} = \exp(\imath\,\mathrm{acos}(H))$. In particular, we can isometrise any generalised translation after rescaling: $T_i^{\mathrm{iso}} = \exp(\imath\,\mathrm{acos}(\frac{1}{\sqrt{N}}T_i) = T_i + \imath\sin(\mathrm{acos}(\frac{1}{\sqrt{N}}T_i))$.

Additionally, the graph translation shares several properties with the time shift. First, spectral components of higher frequencies are more altered than those of lower frequencies: $\|\chi_l - T_{\mathcal{G}}\chi_l\|_2^2 = 2(1-\cos(\omega_l))$. The set of powers of the graph translation is a mathematical group: $T_{\mathcal{G}}^k T_{\mathcal{G}}^l = T_{\mathcal{G}}^{k+l}$. Together with the invertibility property $T_{\mathcal{G}}^{-1} = \exp(+\imath\pi\sqrt{L/\rho_{\mathcal{G}}})$, its action can be reverted: $T_{\mathcal{G}}^{-k}T_{\mathcal{G}}^k X = X$. However, contrary to the time shift, the graph translation is not a convolution by a delta: $T_{\mathcal{G}}^k X \neq \delta_k * X$.

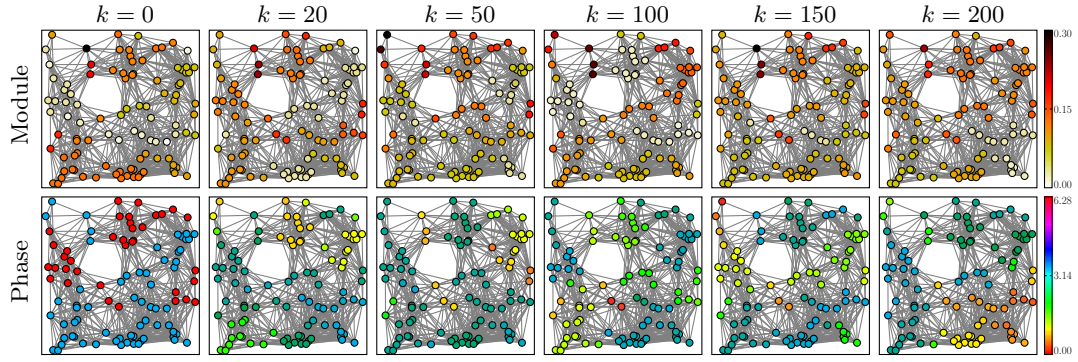
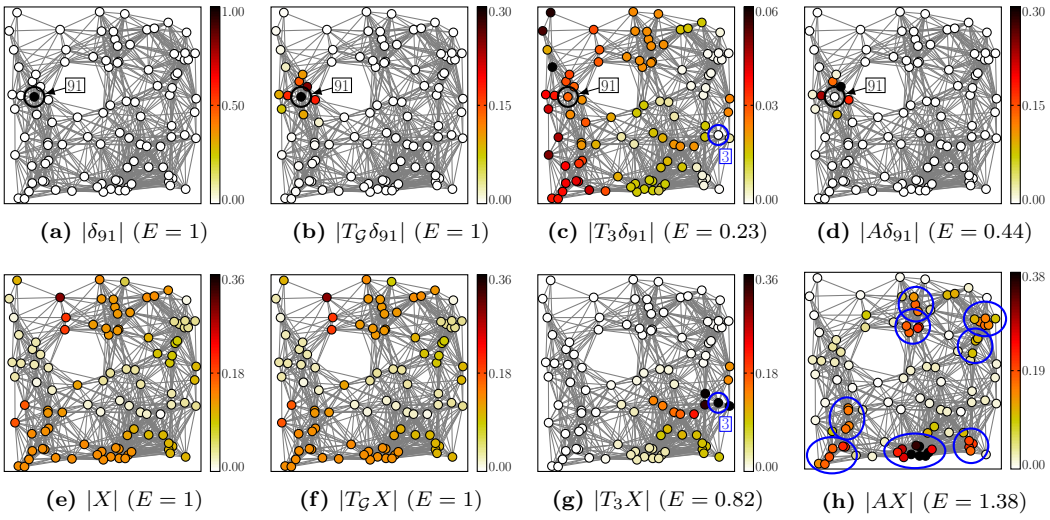


Figure 1. Unit square with 100 vertices uniformly drawn at random. Two vertices are connected if their Euclidean distance is less than $d_{\max} = 0.3$, and weighted by a Gaussian kernel of the distance $e^{-300\,\mathrm{d}(i,j)^2}$. The graph translation operator is iterated on the normalised heat kernel ($\tau = 0$), *i.e.* on the signal X such that $\widehat{X}(\lambda_l) = Ce^{-50\lambda_l}$. The translated signals T^kX being possibly complex, the colour scale illustrates its module on top and its phase on the bottom.



(a) $|\delta_{91}|$ $(E = 1)$ **(b)** $|T_{\mathcal{G}}\delta_{91}|$ $(E = 1)$ **(c)** $|T_3\delta_{91}|$ $(E = 0.23)$ **(d)** $|A\delta_{91}|$ $(E = 0.44)$

(e) $|X|$ $(E = 1)$ **(f)** $|T_{\mathcal{G}}X|$ $(E = 1)$ **(g)** $|T_3X|$ $(E = 0.82)$ **(h)** $|AX|$ $(E = 1.38)$

Figure 2. Same graph structure as Figure 1. Various shift-like operators applied to a delta signal (top row) and the heat kernel X of Figure 1 (bottom row). On (h), the localisation of the lowest frequency components, in the sense of [9], are circled.

4 Graph Translation Illustration

We introduced the graph translation in the Fourier domain as a phase shifting operator on the spectral components. In order to show its behaviour in the vertex domain, we apply our operator to some toy signals. Having a well-defined isometric operator such as the graph translation comes at a cost, namely its complex nature. Its output is then complex for a real input. In this section, we focus on the magnitude of its output. An example of a complex signal phase can be found in Figure 1.

For illustrative purposes, we use a graph whose vertices are drawn uniformly at random in a unit 2D square. We use then the Euclidean distance $d(i, j)$ between vertices i and j to define the weights of the edges through a Gaussian kernel: $a_{ij} = \exp(-\frac{d(i,j)^2}{2\sigma^2})$. Finally, edges between vertices at distance greater than $d_{\max}$ are removed. An example of such a graph is shown in Figure 1 with $\sigma^2 = 1/150$ and $d_{\max} = 0.3$.

Hereafter, we use two toy signals supported by this particular graph to study the behaviour of our operator in the vertex domain. The first signal is a low-band signal. We denote X the normalised signal defined in the spectral domain as $\widehat{X}(l) = Ce^{-\kappa\lambda_l}$. C is such that $\|X\|_2 = 1$. This signal is a *heat kernel* as suggested in [12]. An example of such a heat kernel is shown on Figure 2(e) with $\kappa = 50$. The second signal is a simple delta signal δ_i centred on vertex i. This signal has high frequency components, but more importantly, it allows to characterise the impulse response of an operator.

Figure 1 shows the behaviour of the iterations of the graph translation $T_{\mathcal{G}}$ on the heat kernel X. Transitions between iterations are slow as expected from a low-band signal. Note that the response of $T_{\mathcal{G}}$ to the heat kernel X is very close to X: $\|X - T_{\mathcal{G}}X\|_2 = 0.09$. On the other hand, $T_{\mathcal{G}}$ applied to a delta signal as in Figure 2(b) sensibly alters its input: $\|\delta_{91} - T_{\mathcal{G}}\delta_{91}\|_2 = 1.19$. In both cases, the energy of the signal is preserved along with its PSD. The response of $T_{\mathcal{G}}$ to a delta signal in Figure 2(b) is similar to that of a diffusion operator.

The response of the graph shift A [10] to a delta signal is also that of a diffusion operator (Figure 2(d)). However, as suggested in [9], using the Fourier transform associated to A, the

power spectrum of the signal is changed according to the magnitude of the eigenvalues of A. This can be observed on Figure 2(h) where the lowest frequencies, *i.e.* the highest eigenvalues of A [9], are amplified (circles on Figure 2(h)). Iterating the graph shift further reinforces this bias of the operator towards the low frequencies, as opposed to $T_{\mathcal{G}}$ that preserves the PSD.

Finally, the output of the generalised translation T_3 [12] to X in Figure 2(g) shows a localising behaviour of the operator. However, this is true only for this kind of low-band input signal. For example, its impulse response is hard to read or understand (not shown here). The set of these operators do not form a mathematical group since T_iT_j is not itself a generalised translation [12], when the set of the powers of the graph translation $\{T_{\mathcal{G}}^k\}_k$ is a mathematical group.

5 Stationarity for Graph Signals

Stochastic analysis for temporal signals have produced very interesting results, and we naturally wish to leverage the full potential of the statistics of graph signals to analyse them. Naturally, we need definitions, concepts and associated characterisations for that purpose. As seen in the previous section, stationnarity is one of those important concepts and we propose to define its counterpart for graph signals.

For temporal signals, stationarity can be interpreted as the statistical invariance by time shifting. There exists in the literature two operators defining equivalents of the time shift for graph signals. The generalized translations in [11] define operators acting as generalized convolution by a delta centred on vertex i whose expression in the Fourier domain is $\widehat{T_iX}(l) = \sqrt{N}\widehat{\delta}_i(l)\widehat{X}(l)$. These operators, when applied to very specific input signals, localise their input around vertex i. The authors of [10] proposed the graph shift as the matrix multiplication $X \mapsto AX$. The graph shift is then an operator diffusing a sample from one vertex to its neighbours according to the edge weights.

Unfortunately, neither operator is isometric with respect to the l_2-norm, *i.e.* $\|T_iX\|_2 \neq \|X\|_2$ and $\|AX\|_2 \neq \|X\|_2$. As opposed to the time shift, these two operators lack the mathematical comfort of isometry to define statistical invariance. In another publication, we proposed an alternative time-shift-like operator for graph signals which is isometric by design. After recalling its definition, we use it to propose a definition of stationary graph signals.

5.1 Definition of stationary signals

We now propose to define stationarity for graph signals as the statistical invariance to the graph translation:

Definition 3 (Strict-Sense Stationary)**.** *A stochastic signal* $\mathbf{X}$ *on the graph* $\mathcal{G}$ *is* Strict-Sense Stationary *(SSS) if and only if:*

$$\mathbf{X} \overset{d}{=} T_{\mathcal{G}}\mathbf{X}. \tag{11}$$

Definition 3 is tractable thanks to the isometric nature of the graph translation. However, as for temporal signals, SSS is difficult to verify in practice, and we introduce a weaker, more practical definition of stationarity:

Definition 4 (Wide-Sense Stationary)**.** *A stochastic signal* $\mathbf{X}$ *on the graph* $\mathcal{G}$ *is* Wide-Sense Stationary *(WSS) if and only if:*

$$\mathbb{E}[\mathbf{X}] = \mathbb{E}[T_{\mathcal{G}}\mathbf{X}] \tag{12}$$

$$\mathbb{E}[\mathbf{X}\mathbf{X}^*] = \mathbb{E}[(T_{\mathcal{G}}\mathbf{X})(T_{\mathcal{G}}\mathbf{X})^*]. \tag{13}$$

The invariance property is interpreted as follows: Given a vertex $i \in V$, the random variables $\mathbf{X}_i$ and $(T_\mathcal{G}\mathbf{X})_i$ are statistically equal (SSS) or have identical moments of the first and second orders (WSS). However, in contrast to temporal signals, given two different vertices i and j, there is in general no equality in law between $\mathbf{X}_i$ and $\mathbf{X}_j$.

The definition of WSS signals is introduced in the vertex domain. As for temporal signals, we propose a dual characterisation of WSS signals in the spectral domain. We begin by the first moment. Let $\eta = \mathbb{E}[X]$ be the mean of the graph signal. Equation 12 is then equivalent to $\eta = T_\mathcal{G}\eta$. In other words, η is an eigenvector of $T_\mathcal{G}$ associated to the eigenvalue $1 = e^0 = e^{\omega_0}$. Since $\lambda_1 > 0$, η is collinear to χ_0: The vector of mean is collinear to the Fourier mode of frequency $\omega_0 = 0$. As for temporal signals, the mean of the signal is the equivalent of its DC component.

Next, we characterise the second moment. Let $R = \mathbb{E}[\mathbf{X}\mathbf{X}^*]$ be the autocorrelation matrix of the signal. Equation 13 gives then:

$$R = \mathbb{E}[(T_\mathcal{G}\mathbf{X})(T_\mathcal{G}\mathbf{X})^*] = T_\mathcal{G}\mathbb{E}[\mathbf{X}\mathbf{X}^*]T_\mathcal{G}^* = T_\mathcal{G}RT_\mathcal{G}^*.$$

Let $S = \mathbb{E}[(F\mathbf{X})(F\mathbf{X})^*]$ be the autocorrelation matrix of the Fourier transform of the signal. By linearity of the Fourier transform, we obtain $S = FRF^*$ and $R = F^*SF$. The equality above becomes:

$$S = \widehat{T_\mathcal{G}} S \widehat{T_\mathcal{G}}^*. \tag{14}$$

Assuming that all eigenvalues of L are distinct, then (14) is verified if and only if S is diagonal. In other cases, S diagonal is a sufficient, but not necessary, condition. We can now formally write the spectral characterisation of WSS signals:

Proposition 1 (Second Moment Characterisation)**.** *The second moment of a graph signal is invariant through graph translation if its spectral components are uncorrelated (sufficient condition). This condition is also necessary if all graph frequencies are distinct.*

Proposition 2 (Spectral Characterisation)**.** *A graph signal* $\mathbf{X}$ *is WSS if and only if:*

1. $\mu_\mathbf{X} \propto \chi_0$
2. *if* $\omega_l \neq \omega_k$*, then* $(S_\mathbf{X})_{lk} = 0$

The eigenvalue uniqueness condition is actually verified on many real world graphs. *i.e.* when weights depend on measurements. This comes from the fact that a small random perturbation of weights will slightly change the eigenvalues making them easily unequal. On the contrary, synthetic graphs showing high regularity, such as cycles, or regular grids with unit weights have many multiple eigenvalues [2].

Proposition 1 is similar to the spectral characterisation of WSS time series (*resp.* WSS signals) where the spectral components (*resp.* the spectral increments) are uncorrelated. We have then a doubly orthogonal decomposition.

5.2 Discussion

We remark first that the operator $T_\mathcal{G}$ is in general a complex operator, meaning that if $\mathbf{X}$ is real-valued, then $T_\mathcal{G}\mathbf{X}$ is usually complex. For WSS signals, this is not an issue since Proposition 1 shows that only the correlation between spectral components matters: A signal can be both WSS and real, and both sides of (13) are real for WSS signals.

We consider now the SSS property in the particular case of real-valued signals. Let $\mathbf{X}$ be such a signal. Definition 3 and the Fourier transform give then:

$$\forall i, \forall x, \mathbb{P}[\mathbf{X}_i = x] = \mathbb{P}[(T_\mathcal{G}\mathbf{X})_i = x],$$

and by linear combination of the equalities above:

$$\begin{aligned} & \forall l, \forall \widehat{x}, \mathbb{P}[\widehat{\mathbf{X}}_l = \widehat{x}] = \mathbb{P}[(\widehat{T_{\mathcal{G}}\mathbf{X}})_l = \widehat{x}] \\ \Leftrightarrow \quad & \forall l, \forall \widehat{x}, \mathbb{P}[\widehat{\mathbf{X}}_l = \widehat{x}] = \mathbb{P}[e^{\imath\omega_l}\widehat{\mathbf{X}}_l = \widehat{x}]. \end{aligned}$$

When $l \neq 0$, $e^{\imath\omega_l}\widehat{\mathbf{X}}_l$ is complex and $\widehat{\mathbf{X}}_l$ is real, such that the probabilities above are non zero only for $\widehat{x} = 0 = \widehat{\mathbf{X}}$. Therefore:

$$\forall l \neq 0, \mathbb{P}[\widehat{\mathbf{X}}_l = 0] = 1.$$

A real SSS signal is reduced to a DC component of random amplitude. This case corresponds to the deterministic relation $\mathbf{X} = T_{\mathcal{G}}\mathbf{X}$ since a DC component is invariant through the graph shift.

On the other hand, complex graph signals do not suffer from this constraint. This contrasts with temporal SSS signals that can be real without reducing to a DC component.

Another remark concerns the autocorrelation matrix R. We know that for temporal WSS signals, this matrix is Toeplitz. This is not the case for graph signals. This is to be expected since having such a structure would mean having an autocorrelation function γ which is a function of the difference between vertex indices. In general such a difference does not make sense. For example, shifting all indices by one would leave the matrix R unchanged. However, the graph translation does not perform such a shift such that it is unlikely that the signal is invariant to both the graph translation and this vertex index shifting operation.

Next, the characterisation of the first order moment in the previous section is actually independent of the underlying graph structure. Indeed, χ_0 being always a constant vector, the mean of a WSS signal is constant across vertices. However, several definitions of Fourier transform for graph signals exists, F being one of them. There exists an alternative definition based on the *normalized Laplacian* $\mathcal{L} = D^{-1/2}LD^{-1/2}$ [11]. Both matrices share similar properties. In particular, they are both semi-definite positive. We denote ψ_l the eigenvector associated to the eigenvalue μ_l of $\mathcal{L}$. Proposition 1 is still valid for the Fourier transform $\mathcal{F}$. However, and as opposed to χ_0, ψ_0 is not constant and equals $(\sqrt{d_1}, \ldots, \sqrt{d_N})^T$. We denote $\mathcal{T}_{\mathcal{G}}$ the associated graph translation. We can then use this graph translation to obtain an alternative definition of stationarity. A connected graph verifies $\mu_1 > 0$ such that a WSS signal has a mean vector collinear to ψ_0. Also, SSS real-valued graph signals are now collinear to ψ_0, and not constant. Finally, results on the second moment can be adapted to $\mathcal{T}_{\mathcal{G}}$ directly.

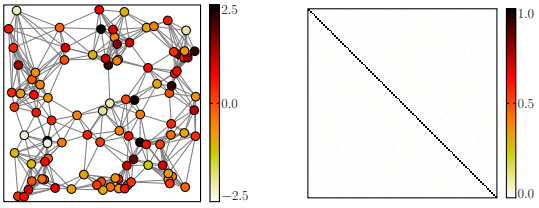
The use of F or $\mathcal{F}$ carries then a different notion of DC component, adapted or not to the local structural properties of the vertices. This accounts for the heterogeneity between vertices weighing differently in the graph structure. Using the Laplacian matrix or the normalized Laplacian matrix carries then different meanings, and using one or the other depends on the application at hand. The rest of this communication will use F as the Fourier transform.

6 Applications of Stationarity

In this section, we wish to apply the notion of stationary graph signals we introduced to concrete graph signals. First, using synthetic signals, and then to study a dataset and the stationarity of a real world signal.

6.1 Synthetic data

In this section, we study several simple stochastic signals. To that end, we use the graph of Figure 3(a). This graph has 100 vertices randomly sampled in the 2 dimensional plane, each



(a) A realisation. **(b)** Empirical S.

Figure 3. White noise on a graph with $\sigma^2 = 1$ and zero mean. (a) shows a realisation of this white noise, and (b) the empirical spectral correlation matrix computed using 50k realisations.

connected to its 8 nearest neighbours. The edge weights are defined by a Gaussian kernel of the Euclidean distance in the plane ($a_{ij} = \exp(-d(p_i, p_j))$).

We propose to define *white noise* on graphs as a stochastic signal having a flat power spectrum, *i.e.* with $S = \sigma^2 I_N$, with mean collinear to χ_0. A realization of white noise with zero mean is shown on Figure 3(a). Using Proposition 1, a white noise is a WSS signal. Since the Fourier matrix F is unitary, we have also $R = \sigma^2 I_N$. Therefore, the samples on vertices are uncorrelated, and of equal variance.

We consider now a signal $\mathbf{X}$ with samples on the vertices independent and identically distributed (i.i.d.). Then, $R = \sigma^2 I_N = S$. $\mathbf{X}$ is therefore a white noise. The whole class of white noise is actually independent of the underlying graph structure. Indeed, the samples being uncorrelated, the edges play no role in explaining any correlation between them.

The third signal we consider is a WSS signal with non-constant spectral power, *i.e.* with S diagonal but of non constant diagonal. Without loss of generality, we suppose the mean of the graph signal to be constant. Figure 4 shows the empirical correlation matrices R and S obtained using 50k realisations. In general, such a signal has correlated samples, but in contrast to temporal signals, the second moment is in general not constant across vertices, *i.e.* the diagonal of R is not constant. Therefore, the mere fact that the signal is WSS is not a guarantee that the samples have the same variance.

This is an illustration of the property that WSS depends on the underlying graph structure. Also, as soon as the signal shows some correlations between vertices, the edges of the graph contribute to explaining those correlations.

6.2 Application on Real Data

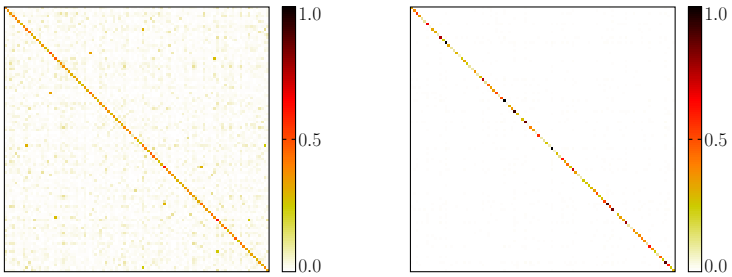
We study now the stationarity on real data supported by a graph to apply the results of section 5, but also to describe the use of empirical estimators on which the characterisation of stationarity relies.

6.2.1 Weather Reports

The dataset studied has been published by Météo France[1] and includes the hourly data reported by weather ground stations in Brittany during the month of January 2014. Overall, this represents 744 readings of 25 ground stations of the strength of the wind, the temperature, and the precipitations.

Each of these ground stations is geolocalised. We use the Euclidean distance to define a graph of the ground stations similarly to section 4, with $\sigma^2 = 5.10^8$ and $d_{\max} = 96$km. This weighting of the edges works well in this context [1].

[1] https://www.data.gouv.fr/



(a) R. **(b)** S.

Figure 4. Empirical second moments of a WSS signal with zero mean and non uniform spectral power with 50k realizations.

The question now is the characterisation of (non)-stationarity for this dataset.

6.2.2 Pre-processing

The stationarity being characterised in Proposition 2 using the statistical moments, these moments need to be estimated using empirical estimators. First of all, using the first empirical moment of the raw data, we can conclude that these data are not stationary. Indeed, the temporal mean of the temperature shown on Figure 5(a) is not constant, therefore not collinear to χ_0. The same phenomenon can be observed for the wind data.

As a consequence, the time series are centred to study the variability of the measurements around the means. Unfortunately, these centred data are not independent from one time instant to another since they are composed, among other components, of seasonal trends. These trends are shown on Figure 5(b). If we are to use empirical estimator (using an hypothesis of ergodicity) of the matrices R and S, we need to ensure the temporal stationarity of the time series.

To remove these trends, we use the *Empirical Mode Decomposition* (EMD) [7]. This decomposition split a time series in several modes from fast oscillations (first *Intrinsic Mode Function* (IMF)) to slow oscillations (last IMF). The first IMF does not yield any temporal trends. The result is shown on the bottom graph of Figure 5(b).

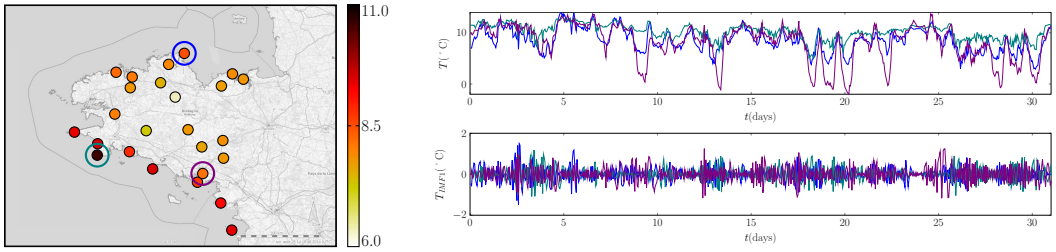
This operation gives good results on both the wind and the temperature. On the other hand, the EMD reaches its limits for precipitations. Indeed, these data show strong saturation behaviours when no precipitation happen. The EMD does not behave correctly in the presence of those saturations. In the rest of this communication, we study the IMF1 of the wind and the temperature.

6.2.3 Study of the Stationnarity

The stationarity being characterised by a diagonal matrix S, it is easier to work with the matrix of spectral correlation coefficients in order to account for the diagonal nature of S. Let $\mathbf{x}$, $\mathbf{y}$ two random variables. Their correlation coefficient is:

$$\mathrm{c}(\mathbf{x},\mathbf{y}) = \mathrm{cov}(\mathbf{x},\mathbf{y})/(\sigma_{\mathbf{x}}\sigma_{\mathbf{y}})$$

with $\mathrm{cov}(\mathbf{x},\mathbf{y})$ their covariance, and $\sigma_{\mathbf{x}}$ the standard deviation of $\mathbf{x}$. We denote C_S the matrix of the correlation coefficients associated with S and such that $(C_S)_{ij} = \mathrm{c}(\widehat{\mathbf{X}}_i, \widehat{\mathbf{X}}_j)$, and C_R the matrix associated with R. Since the IMF1 are centred, their mean is zero, and their spectral components are also of zero mean. Therefore, R or S diagonal is equivalent to C_R or C_S diagonal.



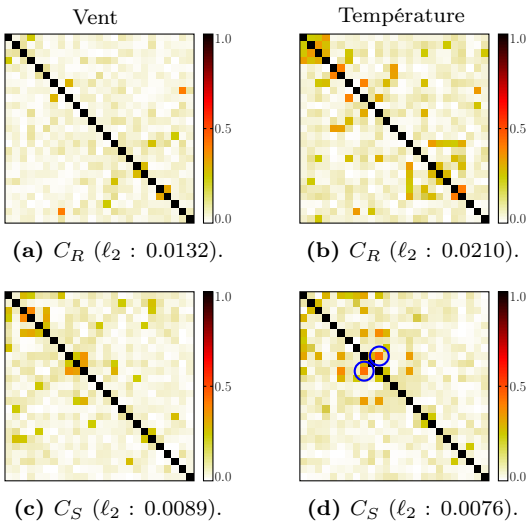
(a) Mean temperature. **(b)** Temporal variations.

Figure 5. Spatial and temporal variations of the temperature in Brittany during the month of January 2014. (a) shows the mean temperature of each of the ground stations, and (b) the temporal variations of three of them of the raw data (top) and of the first IMF (bottom).

The IMF1 of the wind and temperature data give the empirical matrices C_R and C_S shown on Figure 6. We remark on Figure 6(a) and 6(b) that the C_R matrices are not diagonal, illustrating the spatial correlations of the data. This is more visible for the temperature than the wind.

Moreover, it appears on Figure 6(c) and 6(d) that the matrices C_S are diagonal either. We remark also that the stronger correlation coefficient of the temperature corresponds to two Fourier modes co-localised in the South-Ouest region, suggesting that the correlation is a by-product of the structure rather than the data. This correlation is circled in blue on Figure 6(d) and the associated modes are illustrated on Figure 7. Naturally, this structural correlation is not incompatible with intrinsically non-stationary data.

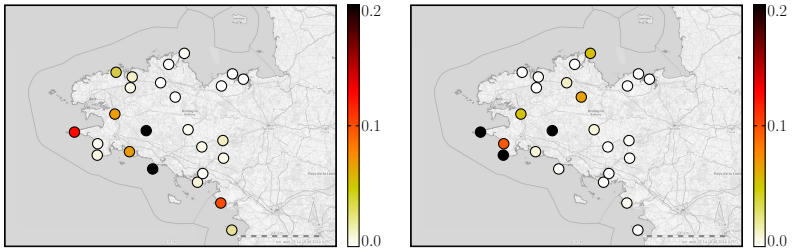
The matrices C_S being non diagonal, the graph signals are not formally stationary. In the next section, we show how to understand this non stationarity are separate the spatial correlations of the measurements from those of the structure.



(a) C_R (ℓ_2 : 0.0132). **(b)** C_R (ℓ_2 : 0.0210).

(c) C_S (ℓ_2 : 0.0089). **(d)** C_S (ℓ_2 : 0.0076).

Figure 6. Matrices of the correlation coefficients in the vertex domain (C_R) and the spectral domain (C_S) for the wind and temperature data.



(a) Mode 9. **(b)** Mode 11.

Figure 7. Localisation of the energy of two Fourier modes of the ground stations graph weighted by a Gaussian kernel of the distances.

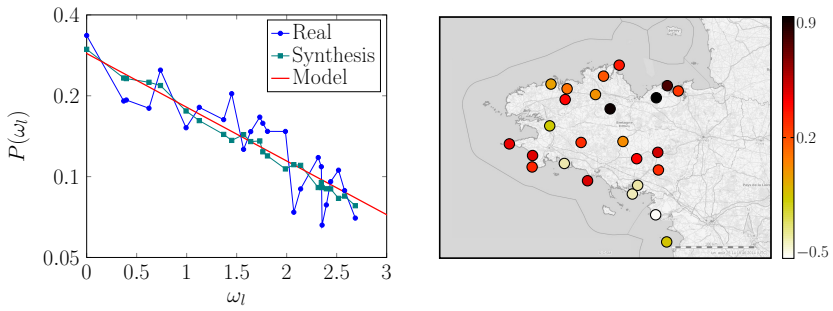
6.2.4 Interpretations

First, we show the influence of geographic structure of the data on the spatial correlations. We consider the temperature data, and the matrices of Figure 6(b) and 6(d). The diagonal of the associated matrix S contains the empirical *Power Spectrum Densities* (PSD) of these data. We observe that the PSD follows a power low plotted in red on Figure 8(a) with the empirical PSD in blue.

In order to study the influence of the spectral correlations on the spatial correlations, we use this model to synthesise realisations of the stationary graph signal of zero mean and of spectral correlation matrix S prescribed by the previous law. We use Gaussian distributions on the vertices as observed on the data.

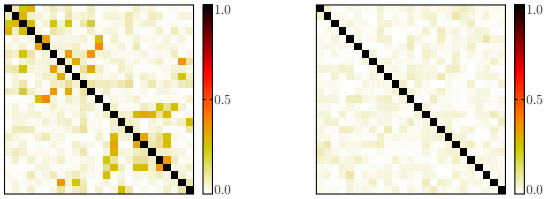
The Gaussian hypothesis allows for a simple generation of the realisations using the Cholesky decomposition $S = LL^*$, with L lower triangular. Let Y be a realisation of N Gaussian variables independent, centred and normalised. Let $\widehat{X} = LY$. The spectral correlation matrix of $\mathbf{X}$ is then S.

Figure 9 shows the empirical matrices C_R and C_S of the synthesised signals. In accordance with the model, the matrix C_S is indeed diagonal. To quantify the differences between these matrices and those of the data, we propose to use the ℓ_2-norm of the non diagonal significant coefficients (*i.e.* the coefficients of p-value less than 0.1), normalised by the number of these coefficients. We observe then a value for this criterion inferior by 20% for the synthesised data: the model does not perfectly describe the data.



(a) PSDs. **(b)** Temperature.

Figure 8. Synthetic data generated from a PSD model in $1/\omega_l^\alpha$ obtain by linear regression of the empirical PSD of the temperature. (a) shows the PSD of the data with the model and the empirical PSD obtained from 744 realisations. (b) shows one of these realisations.

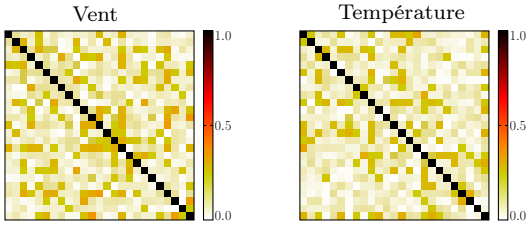


(a) C_R (ℓ_2 : 0.0224). **(b)** C_S (ℓ_2 : 0.0062).

Figure 9. Correlation coefficients matrices of stationary graph signal of prescribed PSD following $1/\omega_l^\alpha$, with as many realisation as Figure 6.

However, the matrices C_R of the data and the model show strong similarities. More precisely, our ℓ_2 criterion applied to the difference between C_R matrices is 0.0065, which is a low value compared that of the C_R matrices alone. Therefore, we can explain a high number of spatial correlations and the PSD using the graph structure we defined.

Conversely, we remove the spatial coherence of the graph structure and study the stationarity of the data on this new graph. To that end, we re-organise the ground stations such that they form a cyclic graph, in an arbitrary order. The edges of this graph are therefore not an image of the geographic organisation of the ground stations. We study now the impact of this re-organisation on the stationarity. Figure 10 shows the matrices C_S of the data on this cyclic graph. We observe an increase of factor 2 of our ℓ_2 criterion compared to the natural graph. These correlations are also more evenly spread within the matrix.



(a) C_S (ℓ_2 : 0.0190). **(b)** C_S (ℓ_2 : 0.0137).

Figure 10. Spectral correlation coefficients matrices of the wind and temperature data using a cyclic graph connecting the ground stations in an arbitrary order.

Moreover, the spatial correlations are completely different compared to the natuarl graph, such that the cyclic organisation of the ground stations does not allow to interpret the observed spatial correlations. Such a graph does not explain the non-stationarity by its structure. Therefore, when deleting the spatial organisation of the vertices, non-stationarity appears to be stronger compared to the use of a more natural spatial organisation.

References

[1] Mikhail Belkin and Partha Niyogi. Towards a theoretical foundation for Laplacian-based manifold methods. *Journal of Computer and System Sciences*, 74(8):1289–1308, 2008.

[2] Andries E. Brouwer and Willem H. Haemers. *Spectra of graphs*. Universitext. Springer, New York, 2012.

[3] Fan R. K. Chung. Lectures on spectral graph theory. *CBMS Lectures, Fresno*, 1996.

[4] Ronald R. Coifman and Mauro Maggioni. Diffusion wavelets. *Applied and Computational Harmonic Analysis*, 21(1):53–94, 2006.

[5] Kinkar Ch. Das. Extremal graph characterization from the bounds of the spectral radius of weighted graphs. *Applied Mathematics and Computation*, 217(18):7420–7426, 2011.

[6] David K. Hammond, Pierre Vandergheynst, and Rémi Gribonval. Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis*, 30(2):129–150, 2011.

[7] Norden E. Huang, Zheng Shen, Steven R. Long, Manli C. Wu, Hsing H. Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, and Henry H. Liu. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *The Royal Society of London. Proceedings. Series A. Mathematical, Physical and Engineering Sciences*, 454(1971):903–995, 1998.

[8] Markus Pueschel and Jose M. F. Moura. Algebraic Signal Processing Theory. Available online: `http://arxiv.org/abs/cs.IT/0612077v1`, December 2006.

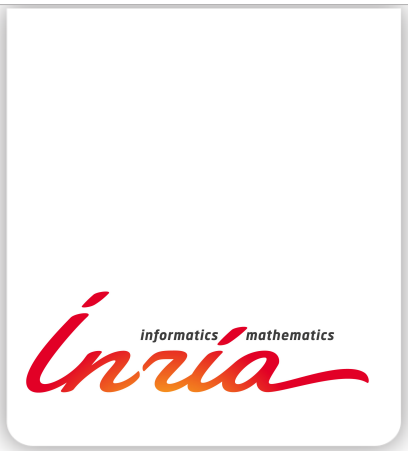
[9] A Sandryhaila and J.M.F. Moura. Discrete Signal Processing on Graphs: Frequency Analysis. *Signal Processing, IEEE Transactions on*, 62(12):3042–3054, June 2014.

[10] Aliaksei Sandryhaila and José M. F. Moura. Discrete Signal Processing on Graphs. *IEEE Transactions on Signal Processing*, 61(7):1644–1656, 2013.

[11] David I. Shuman, Sunil K. Narang, Pascal Frossard, Antonio Ortega, and Pierre Vandergheynst. The Emerging Field of Signal Processing on Graphs: Extending High-Dimensional Data Analysis to Networks and Other Irregular Domains. *IEEE Signal Processing Magazine*, 30(3):83–98, 2013.

[12] David I Shuman, Benjamin Ricaud, and Pierre Vandergheynst. Vertex-frequency analysis on graphs . *Applied and Computational Harmonic Analysis*, 2015. in press.

[13] D.I. Shuman, B. Ricaud, and P. Vandergheynst. A windowed graph Fourier transform. In *Statistical Signal Processing Workshop (SSP), 2012 IEEE*, pages 133–136. IEEE, 2012.



RESEARCH CENTRE
GRENOBLE – RHÔNE-ALPES

Inovallée
655 avenue de l'Europe Montbonnot
38334 Saint Ismier Cedex

Publisher
Inria
Domaine de Voluceau - Rocquencourt
BP 105 - 78153 Le Chesnay Cedex
inria.fr

ISSN 0249-6399